

Do Donors Get What They Paid For?

Micro Evidence on the Fungibility of Development Project Aid

Dominique van de Walle and Dorothyjean Cratty*

World Bank

Abstract

Recipient government responses to development project aid have typically been studied at high levels of aggregation, using cross-country comparisons and/or aggregate time series data. Yet increasingly the relevant decisions are being made at the local level, in response to specific community-level projects. This paper uses local-level survey data collected for this purpose to test for fungibility of World Bank financing of a rural road rehabilitation project targeted to specific geographic areas of Vietnam. A simple double difference estimator suggests that the project's net contribution to rehabilitated road increments is close to zero, implying complete displacement of funding. However, with better controls for the endogeneity of project placement we find much less evidence of fungibility, with displacement accounting for around one-third of the aid. The results point to the importance of dealing with selection bias in assessing project aid fungibility.

JEL classification: O19; O12; F35;

Key words: Impact evaluation, fungibility of aid, rural roads, Vietnam

World Bank Policy Research Working Paper 3542, March 2005

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the view of the World Bank, its Executive Directors, or the countries they represent. Policy Research Working Papers are available online at <http://econ.worldbank.org>.

* We thank Vu Tuan Anh, Christina Malmberg-Calvo, Binh Nguyen and Martin Ravallion for their encouragement, help and support. We received useful comments on this paper from Kathleen Beegle, Michael Clemens, Hai-Anh Dang, Jishnu Das, Quy-Toan Do, John Giles, Martin Ravallion and participants at the 2004 NEUDC meetings in Montreal. We gratefully acknowledge funding support from CIDA (TF034859), the World Bank's East Asia and Pacific Transport Unit and the Poverty and Social Unit in the Vietnam Resident Mission Office, and the DFID trust fund for Poverty Analysis and Policy Advice. Please do not cite without authors' permission. Address for correspondence: Dominique van de Walle, World Bank, 1818 H ST, NW, Washington, DC 20433. tel: 202-473-7935, fax: 202-522 1154, dvandewalle@worldbank.org.

1. Introduction

It is well recognized that the degree of fungibility of external aid has important implications for assessing aid effectiveness. When aid that is ostensibly tied to a specific project intervention is in fact fungible, it becomes extremely difficult to determine what exactly the aid has financed and hence what development impact it has had on the ground.

Empirical investigations of development aid fungibility have tended to be based on cross-country regressions or assessments of macro aggregates over time. On balance, most empirical studies find evidence of substantial fungibility, despite earmarking. Yet development assistance continues to be primarily project-based, particularly for infrastructure. There appears to be little or no evidence based on actual development projects, which are often implemented at the local level, under increasingly more decentralized fiscal arrangements. The aggregate data used in past work may have little power to detect fungibility in such settings.

At the same time, the issue of fungibility has been routinely ignored in project work. Project staff often spend an inordinate amount of energy on how best to target and how to take social issues into account in project selection and appraisal. If fungibility is indeed the reality, then all these efforts are moot. This is not to deny the importance of these concerns but to question whether their emphasis and discussion should not be de-linked from projects. To take just one current example, UNDP's recent millennium report calls for funding to be directed to priorities identified in a needs assessment for Africa (UNDP, 2004). A per capita amount is assessed for a number of key areas. Fungibility is not considered. Yet, it may make a considerable difference to the eventual outcomes of

such a policy and to the cost calculations on which the policy choice is based.

This paper studies the degree to which road project funding actually ends up funding what the donor intended. We focus on the first World Bank-financed Rural Transport Project (RTP1) in Vietnam — a project launched in 1997 and aiming to rehabilitate 5,000 kilometers of district and commune level roads in 18 provinces.¹ Rural roads are being extensively championed as poverty alleviation instruments by the World Bank and donor institutions. It is argued that rural roads are key to raising living standards in poor rural areas (for example see Gannon and Liu, 1997) and the international financial institutions are scrambling to implement rural road projects. However, rigorous impact assessments have been rare. Such assessments require appropriate controls, and allowing for unobserved factors that influence both program placement and outcomes. This study uses a panel data set (collected for the purpose of conducting such an impact evaluation) of communes and households within both project and non-project areas; the data comprise a (pre-intervention) baseline and two post-project follow-up rounds of data. We use these data to test for fungibility.

Much to our initial surprise, a casual look at the project evaluation data revealed only a slight difference between project and non-project comparison communes in the length of roads undergoing rehabilitation during the project period. A simple mean difference between participating and non-participating communes showed a mere 0.08 kilometers cumulative average difference in favor of project communes over the 5 years. Yet, independent administrative data on the project reported that an average of over 4 kilometers per commune were rehabilitated under the project. Project supervision and

¹ The Vietnam Rural Transport Project I, see World Bank (1996) for details.

other evidence confirm that the project was implemented as planned.

There are at least two possible explanations for this observation. One is fungibility. By this explanation, the recipient government has its own preferences over spending, and so it cuts its own funding of roads in the project areas in response to the external aid. Suppose, for example, that the areas targeted for the project are initially spending \$80 per capita per year on road rehabilitation and those not targeted spend \$100. The project ostensibly allocates an extra \$50 to the former areas. However, the government treats this as an increase in its total budget for roads and decides to spend \$105 in the project areas and \$125 in the other areas. In other words, it cuts its own spending in the project areas by \$25, which is reallocated to the non-project areas. Note that in this hypothetical case the development aid is still having an impact, it is just not identifiable by comparing the increment to road spending in project and non-project areas. However, in the extreme case of fungibility, all external funding for roads goes into consolidated spending and there is no earmarking for roads per se. One then expects an income effect of the extra aid on road (and other) spending, but no differential impact on project areas versus the non-project areas used for comparison purposes.

A second possible explanation for the observed lack of difference in rehabilitated road length is that the non-project communes are not an appropriate comparison group and so our counterfactual is inappropriate. In judging the impact of the project we may need to deal with the possibility of selection bias. The project may have been targeted to places with a worse state of roads but also with less economic potential. A selection bias would occur if there is some latent reason for why roads are poor in participating communes and they are being compared with places that don't have these factors. The

implication of this second explanation is that what looks like fungibility of project aid may in fact be the outcome of biases in the evaluation methods. The principal finding of this paper is that the second explanation is the right one. Failure to deal with the heterogeneity in factors influencing the placement of local development projects can seriously bias assessments of the extent of fungibility.

The paper first reviews the literature on aid fungibility. This is followed by a brief description of the project being evaluated and the rural Vietnam country setting. Section 4 describes the survey instrument, while section 5 discusses the methods used for evaluating the degree of fungibility. Section 6 discusses the results and various sensitivity tests. Section 7 concludes.

2. Development aid fungibility

A number of studies have investigated the degree to which donors' earmarked development or project aid substitutes for rather than supplements local spending intended for that purpose. Fungibility — whereby categorical aid substitutes for government funding by being diverted to spending on other sectors or being used to reduce taxes — tends to be difficult to detect and to identify empirically. The literature has attempted to identify government responses from the variance in spending across countries or from within country changes in spending over time. For example, many past studies have used pooled cross-country time series data to examine the impact of tied aid on government spending across sectors (for example, see Feyzioglu et al., 1998; World Bank, 1998). Another set of studies uses country specific time series of macro aggregates to explore the effect of categorical aid on the public sector's expenditure mix (Pack and Pack, 1990 and 1993). Government responses to targeted aid are found to vary quite a

lot. Nonetheless, although there are some notable exceptions, a significant amount of fungibility tends to be found.²

The approach adopted in this paper to examine the question of project aid fungibility differs from the past literature in a number of fundamental respects. The first major difference concerns the level of aggregation. At the level of aggregation at which such studies are typically conducted, a lot of the variance in project aid and responses to it will be lost. For one, the situation in the absence of aid is not observed. This undoubtedly leads to a loss in the ability to identify the key parameters of interest. In contrast, since external aid projects can't be implemented everywhere, there will be places that don't get them. Observations on places without the project aid provides us with useful extra information that other studies have not had.

A second difference concerns the degree of heterogeneity that can be allowed for in our approach. Differences in government preferences will be difficult to take into account with cross-country regressions. One has a better hope of identifying fungibility using country specific data. But country specific studies have also tended to be highly aggregated and focus only on central government behavior. A finding of no fungibility at the central level could well be compatible with widespread displacement at lower government levels. Our approach allows us to explore whether fungibility goes deeper into local government behavior — at the level where the key decisions are made. Against that, our approach has the drawback that we may be unable to detect fungibility that occurs at the highest level of government.

In the case of the rural road project studied here, diversion could occur either at

² One exception is for Indonesia where Pack and Pack (1990) find no diversion of earmarked aid. For some sectors, foreign aid is even found to have stimulated additional government spending.

the local or higher levels. The central government may well reduce its budget allocation to the Ministry of Transport as a response to the World Bank's project funding. We are unable to detect fungibility at this level and can only test for fungibility at lower levels of government. However, our initial finding of no differential impact of the project at the local level suggests that this is an appropriate level at which to examine the issue.

Funding from intermediate levels of government (district or province) that had been slated for road improvements in the project communes may get displaced to communes not benefiting from the donor financed project. We think this is the most likely way that fungibility occurs in this setting. The local commune authorities may also decide to divert resources that would have been spent on roads to other commune-level needs or to reduce the commune-level tax burden. However, given that communes typically do not have the resources necessary to undertake road rehabilitation, and that under the project, provinces retained substantial control of project funding, displacement is unlikely to have been significant at the commune level. Road work at the commune level — consisting primarily of routine maintenance — is typically implemented by commune residents in Vietnam. External road rehabilitation funding could result in a reduction of residents' 'corvée' labor burden or its redirection to other needs. Either way, there will be less impact of the project on the roads actually rehabilitated in project areas than expected.

Behavioral responses by aid recipients could also mean that road spending is used differently in project areas. For example, when labor inputs are supplied locally and are fixed, project funds may then encourage a switch in local road rehabilitation efforts away from more labor intensive techniques (more sealed roads than unsealed roads, for example). The outcome may be fewer rehabilitated kilometers but higher quality

rehabilitation.

Injecting foreign aid may also lead to general equilibrium effects within an economy so that it becomes difficult to isolate the effects of interest. This also speaks to looking at more local interventions that don't change the entire structure of the economy and have much more localized impacts.

It should be noted that even with full funding fungibility one may still see a positive impact on road spending. Fungibility means that you're free to spend on whatever you like. This will result in a pure income effect on the demand for all normal goods including roads. This poses a problem for measuring fungibility with our approach (as well as with past approaches; see for example the discussion in Khilji and Zampelli, 1994). Because of a potential income effect of the fungible component of aid funding, we cannot actually identify the amount of fungibility. The combination of the two, which we do observe, can be called "displacement." In order to properly isolate fungibility from the income effect we would need an evaluation design whereby some communes were allocated an untied allocation of the same amount of money as received under RTP1. One could then evaluate the effect of that income allocation on road spending in those communes and add that amount to the displaced spending to get at the degree of actual fungibility. Unfortunately we do not have such an evaluation design. Under these circumstances, the potential presence of an income effect must be considered when interpreting our results.

3. The project intervention and selection of participants

Vietnam has poor physical infrastructure and high levels of income poverty (van

de Walle 1998). The country only recently came out of a long period as a planned market economy in the shadow of the Soviet Union when large capital projects were favored and rural infrastructure was long neglected. Since it began to adopt a market economy in 1987, Vietnam has been experiencing remarkable changes in all aspects of economic life, and since the U.S. trade embargo was lifted in 1994, the resumption of bilateral and multilateral aid. The last decade has seen considerable—though geographically unequal—growth. Many observers have argued that basic infrastructure investments, and rural roads specifically, will reduce poverty in Vietnam. This view has been popular with donors as well as with the government.

It is within this larger context that the World Bank set up its first rural transport project in 1997. The Vietnam Rural Transport Project 1 is a World Bank financed large-scale rural roads rehabilitation project aiming to link commune centers to markets and reduce poverty (World Bank 1996). Communes are the lowest administrative unit in Vietnam, immediately superceded by districts and above that provinces. The commune center is typically where the commune's primary school, health center, main economic center and political and administrative facilities are located. The 200 communes in our sample had an average of 1,308 households living in 10 or so hamlets (or villages).

The project was launched in 1997 for implementation in 18 poor provinces over 5 years, at a cost of about \$61 million. It aimed to rehabilitate 5,000 kilometers of roads in total, comprising of 3,500 and 1,500 of district and communal roads, respectively. In each participating province, road links were identified for rehabilitation through least cost techniques.³ In practice, a complete rehabilitation standard was generally enforced

³ Least cost techniques refer in this project to the minimum cost engineering solution that ensures a certain level of motorized passability.

aiming to provide ‘reliable access’ — defined as relatively consistent and safe access with only short-term road closures (due to bad weather). The project stipulated that no new roads should be built. However, the eligible existing roads were in many cases unusable with whole sections that were impassable by a motor vehicle for much of the year.

As noted, RTP1 funds were earmarked for the rehabilitation of commune and district level roads, also denoted rural roads. Spending on rural roads can come from various sources and levels of government. These roads are normally the responsibility of commune and district governments who typically lack the necessary resources for anything other than routine maintenance. In circumstances where other funding such as from the central government or international donors does become available, the provincial authorities are supposed to channel the funds to the targeted communes. The World Bank project is implemented through the central government’s Ministry of Transport (MOT), which was responsible for deciding which provinces would participate in the project. Its provincial offices were then in charge of picking communes and road links to be rehabilitated, and choosing appropriate contractors following a competitive bidding process. The road link selection is supposed to be based on a number of conditions imposed by the project; a proposed road was eligible subject to average investment costs being no more than \$15,000 per km and the population served being at least 300 people per kilometer. In addition, in an effort to extend project benefits to low density, mountainous areas with concentrations of ethnic minority populations, 20 percent of each province’s rehabilitation funds could be set aside for roads not justified under the population and cost criteria. Presumably, the need for rehabilitation will also affect the choice of road links to be rehabilitated by the project.

It should be noted that RTP1 is implemented against a backdrop of large increases in the road network and infrastructure development throughout the country during the 1990s. This points to the importance of not relying on reflexive comparisons (tracking gains solely in project communes) but introducing instead a comparison group.

4. The SIRR data

We use a data set that was created specifically for evaluating the impact on living standards of the rural roads rehabilitated under RTP1. The "Survey of Impacts of Rural Roads in Vietnam" (SIRR) is a panel data set of pre-project baseline and post-project data for both project ("treatment") and non-project ("comparison") areas.⁴ The data were collected in six out of the 18 provinces that are included in RTP1. Two provinces were randomly picked from the participating provinces in the country's north, center and south: Lao Cai and Thai Nguyen in the north, Nghe An and Binh Thuan in the center and Kon Tum and Tra Vinh in the south.

In each of the six provinces, samples of project and non-project communes were drawn for a total of 200 surveyed communes. Road links covered by the project all pass through communes, and a majority link up commune centers—where facilities and services are located—with the road network. Data are often, and more easily, collected at the commune level in Vietnam. For these reasons, the zone of a road's influence is defined as the commune through which the road passes. The project communes were randomly selected from lists of all communes with proposed projects in each province. A list was then drawn up of all remaining communes in districts with proposed sub-projects

⁴ For a more detailed description of the SIRR see van de Walle (2005).

from which a random sample of non-project communes was drawn. The comparison areas chosen this way should share many of the same characteristics as the project areas. However, we cannot be confident that they are a good comparison group on a priori grounds. For this reason, we use matching techniques based on the data collected for the evaluation to test the selection of comparison groups.

A detailed commune-level data base was created in part by drawing on annually collected records—both current and retrospective—at the commune level and augmenting this information with various other supplementary data. This approach takes advantage of Vietnam’s tradition of data collection at the local level. The commune questionnaire includes sections on general commune characteristics, infrastructure, employment, sources of livelihood, agriculture, land and other assets, living conditions, education, health care, development programs, community activities and organizations, commune finance and prices.

In addition, a household questionnaire was administered to 15 households selected through a system of stratified sampling in each sampled commune. Five households were chosen from each of three lists, containing the poorest, middle and richest thirds of all households in the commune. The lists were based on a welfare ranking done by the commune authorities. Clearly, these rankings are to some extent subjective, but stratified sampling on this basis should assure a sample that is reasonably representative of each commune’s main socio-economic groups.

Households were asked about general characteristics, employment, assets and amenities, production and employment activities, participation in and access to education, health, markets, credit, community activities, social security and poverty programs, and

transport. The commune and household questionnaires are primarily quantitative although both also include some qualitative questions.

No attempt is made to measure a household level indicator of welfare such as income or consumption. This decision reflects a careful weighing of the constraints faced by the survey instrument and the severe difficulties involved in collecting reliable welfare information. However, the household questionnaire was designed in view of combining it with the Vietnam Living Standards Survey (VLSS) (Vietnam's only nationally representative household survey) of 1998 to predict consumption expenditures for SIRR V households. As described in van de Walle (2004), we use regression techniques to combine information on household characteristics common to both surveys and estimate baseline consumption expenditures for each SIRR V household.

The baseline data collection began in June 1997. Subsequent rounds followed at two year intervals, in the summers of 1999, 2001 and 2003. (We do not use the last round as it covers the post-project period, which is not a concern here.) Each SIRR V round tracks the implementation process followed for previous rounds of data collection as closely as possible. Communes and households surveyed are the same in each round, though there is some attrition of households. Surveying begins simultaneously in each province, and coverage of districts is timed across months to coincide as much as possible with the schedule followed during the baseline.

We also make use of a project level database that was constructed as part of the research and covers the projects in our survey areas. This provides information on timing, costs, the magnitude of the change resulting from a project as well as the method of project implementation and road kilometers rehabilitated.

Much of the analysis in this paper focuses on a section of the commune questionnaire that asks about the kilometers of roads that were rehabilitated and built during the last two years, as well as details on the related amounts of funding received from different sources. The latter could include the central government, the province, district, or commune budgets, international agencies, and villagers' contributions — consisting of the locally valued labor time contributed.

A potential data problem should be noted. Routine maintenance appears to have been recorded as rehabilitation for some communes.⁵ We are particularly suspicious that these are cases of maintenance when the kilometers rehabilitated equal, or very nearly equal, the commune's entire road or earth road network, the costs are low and the work has been entirely implemented by villagers' labor contributions, with added funding from the commune only, and sporadically from the district's budget. The incidence of such cases is more common in the baseline and diminishes over time. The latter could indicate that road maintenance is actually declining over this period as more funds get allocated to rehabilitation. Or it could reflect a change in the question's interpretation by the interviewers over time. We do not know which it is.

There are two ways in which this potential data problem could matter to the fungibility issue at hand. It may affect selection into the project, but equally well, it may affect the impacts from the project. It seems plausible that the demand for rehabilitation will be low in places where maintenance is well done. If so, the project will tend to be allocated to areas with a poor record in maintenance. One can also imagine that the

⁵ Routine maintenance refers to work that is not intended to cover major repairs or upgrades, and is required annually or periodically on all roads in good and fair condition, rendering rehabilitation unnecessary. Rehabilitation is done on poor and very poor roads that have not received proper maintenance.

propensity to maintain roads is a function of local governance. Hence, poor local governance could affect selection over time. Good governance may also be required to derive impacts from the project. In this story, road rehabilitation efforts are targeted to places with poor administrations of road maintenance and poor governance that may in turn result in low impact. Comparison groups are then likely to have better governance unless we can allow for this selection bias in the matching.

We will test the sensitivity of our results to this possibility. We create a dummy to identify communes that report implementing rehabilitation that looks to us to be routine maintenance in the two years prior to each survey round. The dummy is defined in the following way. We identify all probable maintenance cases as those where the recorded number of kilometers rehabilitated equal either the total number of road, or earth road kilometers in the commune. From among these probable maintenance communes, we then identify those for whom the log of funding per kilometer rehabilitated was 2 standard errors lower than the survey mean funding per kilometer rehabilitated. Unfortunately, the surveys did not ask about routine road maintenance.

5. Evaluation methodology

Despite the non-project communes being originally selected to be “similar” to the project communes, one can expect to find differences between these two sets of communes. For example, in our particular case, this could come from the need for rehabilitation. One would expect the extent of rehabilitation that is implemented to be highly related to initial road infrastructure conditions, raising an endogeneity concern. We would then want to control for initial conditions in terms of the initial state of

infrastructure in choosing comparison communes.

To do so we combine double differencing with propensity score matching (PSM) to select ideal comparison communes from among the sampled non-project communes. The impact of the road infrastructure is then identified by the difference between outcomes in the project areas after the project and before it, minus the corresponding outcome difference in the matched comparison areas. This “matched double difference” estimate will give an unbiased estimate of project impacts in the presence of unobserved time invariant factors influencing both the selection of project areas and outcomes.

PSM finds a non-project commune that is similar in observed covariates to the project commune. This is done by comparing the propensity scores – the predicted probability of a commune getting a project – of project and non-project communes (Rosenbaum and Rubin, 1983, 1985). A non-project commune whose propensity score is the “closest” to or judged to be within an acceptable maximum distance, or caliper, from the propensity scores of a project commune is declared to be a matched comparison commune for that project. Once the matched comparison group is constructed, the mean impact estimate is the difference between the project (treatment) and the matched comparison communes in the different outcome indicators of interest.

There are communes that have a road project ($D_i=1$) and communes that do not ($D_i=0$). The former “treated group” are matched to the “comparison group” on the basis of their propensity scores. The population propensity score for commune i is defined as:

$$P(X_i) = Prob(D_i = 1 | X_i) \quad (0 < P(X_i) < 1)$$

where X_i is a vector of pre-project explanatory variables. Rosenbaum and Rubin (1983) show that if (i) the D_i ’s are independent over all i , and (ii) outcomes are independent of

project participation given X_i (i.e. unobserved differences across the treated and the comparison groups do not influence being in a specific group), then outcomes are also independent of project participation given $P(X_i)$, just as they would be if participation is assigned randomly.⁶ PSM uses $P(X)$ or a monotone function to select comparison subjects for each of those treated. Exact matching on $P(X)$ implies that the resulting matched comparison and treated subjects have the same distribution of the covariates. PSM thus eliminates bias in estimated treatment effects due to observed heterogeneity.

We follow the common practice in PSM applications of using the predicted values from a standard logit model to construct the propensity score for each observation in the treatment and the non-project samples.⁷ Matched-pairs of communes are then constructed on the basis of how close the estimated scores are across the two samples. The nearest neighbor to the i 'th project participant is defined as the non-project that minimizes $[P(X_i) - P(X_j)]^2$ over all j in the set of non-projects. Matches are only accepted if $[P(X_i) - P(X_j)]^2$ is less than a caliper of 0.03.⁸

Beyond the precision of the individual matches, it is necessary to insure that the overall distribution of observables controlled for in the logit is similar for the two groups. Specifically, bias due to unobservables is seen to be significantly reduced by restricting the sample to the common support (Heckman et al., 1997). The final sample of treated observations are only those for which the propensity score is within the maximum propensity score for the control group plus or minus the value of the caliper. This raises a

⁶ Assumption (ii) is sometimes referred to in the literature as the “conditional independence” assumption, and sometimes as “strong ignorability.”

⁷ Dehejia and Wahba (1999) report that their PSM results are robust to alternative estimators and alternative specifications for the logit regression.

⁸ We experimented with more stringent tolerance limits and the results were robust. However, with more stringent limits we also had to discard more participants while calculating our impacts.

concern for selection bias however if the trimming results in a large loss of treated observations with high probabilities of participation. An alternative is to trim only the comparison communes with propensity scores outside the common support — or below the lowest treatment’s propensity score. We will refer to this as outer support matching.

In calculating the average outcome indicator of the matched non-participants several weighting schemes can be used, ranging from “nearest neighbor” weights to non-parametric weights based on kernel functions of the differences in scores (Heckman et al., 1997).⁹ We use the nearest neighbor estimator, which takes the outcome measure of the closest matched non-participant as the counter-factual for each participant.¹⁰

The availability of pre- and post-project data for the same communes allows us to implement a panel data difference-in-difference matching estimator. Hence, we take the difference in the impact estimators across the two periods and across the matched project and non-project communes. An advantage of panel data is that it allows us to control for idiosyncratic unobservables that may influence selection into the program. Assuming that such unobservables can be represented as a time-invariant error component, selection bias can be eliminated by taking first differences over time. Propensity score matching using panel data also allows us to separate the impact of the project from the general road infrastructure investments that would have happened had there been no RTP1 project.

Let Y_{it}^k be the outcome observed for commune i in time $t=(0,1)$ with participation status k (taking the value 1 for participants and 0 otherwise). The difference in mean outcomes between the n treatment communes and the n matched comparison communes

⁹ Jalan and Ravallion (2003) discuss the choice further, and find that their results for estimating income gains from an anti-poverty program are reasonably robust to the choice.

¹⁰ Rubin and Thomas (2000) use simulations to compare the bias in using the nearest five neighbors to just the nearest neighbor; no clear pattern emerges.

in the baseline is

$$\sum_{i=1}^n [Y_{i0}^1 - Y_{i0}^0] / n \quad (1)$$

where Y_{i0}^0 is the value of the outcome indicator for the nearest neighbor to treatment unit i among the non-participants at date 0. The corresponding difference found in the follow-up survey is:

$$\sum_{i=1}^n [Y_{i1}^1 - Y_{i1}^0] / n \quad (2)$$

in obvious notation. Then the difference-in-difference (DD) matching estimator is simply the difference between (2) and (1), namely:

$$DD = \left\{ \sum_{i=1}^n [Y_{i1}^1 - Y_{i1}^0] - \sum_{i=1}^n [Y_{i0}^1 - Y_{i0}^0] \right\} / n \quad (3)$$

6. Results

Tables 1 and 2 present simple descriptive statistics on baseline distances and road characteristics for the project and non-project commune samples. This helps give some feeling for the setting, and also allows a judgment of how similar project and non-project communes appear to be on average, at least with respect to their road and other infrastructure situation and endowments.

Table 1 indicates little difference with respect to average distances in kilometers from their commune centers to a national or provincial road, the closest big city, railway or waterway. The only statistically significant differences are in how far the closest provincial center is (54 versus 44 km for non-project versus project communes) and the distance to the closest provincial road which is also lower for project communes.

Turning to road endowments in Table 2, a larger percentage of project communes have a

national road passing through them. With this one exception, we again find no statistically significant differences in terms of total kilometers of communal roads, and kilometers of communal roads built or rehabilitated during the three years prior to the baseline survey. On the face of it, the two samples appear to be very similar in their access and transport infrastructure characteristics.

Table 3 shows means of kilometers built and rehabilitated during the two years prior to each survey round in the project and non-project communes as reported by the communes. Stars indicate a statistically significant difference between the two groups. The project and non-project commune means follow each other rather closely. This is especially pronounced in the two years prior to the 2001 survey round. Although the values are slightly higher for the project communes starting in the initial situation, none of these differences are significant. Strikingly, the only significant difference is not for kilometers rehabilitated, but for kilometers built, although only in 1999 — which suggests that more road kilometers were built in the project communes between the start of the project and 1999.

6.1 *Matching methods*

Table 4 presents the logit model of commune participation. The dependent variable takes the value one for communes that had a project and zero for those that did not. The aim here is to control for all possible variables that could affect participation in the project and are exogenous to the project. The performance of the PSM technique rests on how well one can predict the probability of participation by different communes in the road rehabilitation intervention. A good predictive model helps justify the key conditional independence (“strong ignorability”) assumption in PSM.

We control for as many initial period commune level characteristics as are available in our data and are not afflicted with missing observations. These include terrain and location, population and population density, the share of Kinh ethnicity households (the majority ethnic group in Vietnam), the school enrollment rate, whether the commune has a production organization, an Agricultural Bank, other sources of formal credit, a market, mineral exploitation, an active land market, and unskilled non-agricultural employment opportunities. A commune's location relative to major social services and commercial centers is identified by the distances to the closest big city, and the province and district centers. Important controls for initial transport infrastructure are road density (of commune and district level roads), the share of total district and commune road kilometers that are impassable earth roads, and whether a national or provincial road, a waterway or railroad pass through the commune. A few measures of commune population characteristics are created from the household data, namely the share of households with locally born heads, with a pensioner, with wage income, with income from cottage industries or from the trade and services sectors, who own a motorcycle, live in a permanent dwelling, and who are landless. Also calculated from the household data is average commune household expenditures.¹¹ Finally, we also control for our measure of whether the commune carried out routine maintenance that it identified as rehabilitation as explained in section 4. Simple tabulations suggest that such communes were no more or less likely to participate in the project. The estimated coefficient is also insignificant.

A key explanatory factor is likely to be the pre-project condition of the road. This

¹¹ This is calculated from the household level estimates based on a consumption model calibrated to the 1998 VLSS and used to predict consumption for SIRR household in 1997.

can be expected to be a key factor both in explaining why a road was chosen for rehabilitation and subsequent impacts, that one would expect to be a function of the difference rehabilitation has made to the road's condition and passability. Our measures of rural road density and the share of impassable earth roads should get at this indirectly. On the argument that if a commune has recently rehabilitated part of its road network, or received funding from the central or provincial governments for this purpose in recent years, it may be less likely to have been chosen to participate in RTP1, one might also want to include measures of these attributes. There is clearly a sense in which “everyone gets their turn” in Vietnam and so we think this may be a powerful explanatory factor for project participation.

However, this raises an endogeneity concern. It seems plausible that the authorities will link their decisions across years — following a sequential or dynamic process whereby a decision is made at the beginning of two periods that any commune that does not receive rehabilitation funding this year will do so next time such funding becomes available. Then both decisions will be endogenous. Matching on variables that are jointly determined with project participation will lead to mismatching. It seems likely that this plays a role in how provinces choose participants. For this reason, we do not include variables such as the percent of road kilometers rehabilitated, or whether the commune received funding from the province or central government levels in the three years prior to the baseline survey even though this information is available in our data. The same argument applies to any other policy variable that is possibly jointly determined by the province or higher levels of government who are also responsible for picking RTP1 participants.

The logit contains no direct measures of political power or bureaucratic efficiency/dynamism at the commune level. It could be argued that such unobservable attributes may have influenced commune selection and will also affect project impacts. Against that, any differences in local government efficiency or political influence are likely to be reflected in initial commune characteristics which the logit does control for. So we are indirectly controlling for such potential differences across communes. And although these same unobservables may matter to other commune level investment decisions, they are unlikely to affect the project's impact on road kilometers rehabilitated. Once allocated to communes by the provincial management units (PMU) within the provincial transport offices, these road link projects are contracted out to bidding companies who are also chosen by the PMUs to whom they also report. Thus the commune authorities neither control nor supervise the road works. Nonetheless, it is of course possible that there is an unobservable factor that influences selection as well as the trajectory of the project impact on kilometers rehabilitated over time. Selection bias on unobservables can never be fully ruled out in observational studies.

A number of explanatory variables are significant. Participation is more likely in communes with higher total population (controlling for density) — a project selection criteria. On the other hand, there is no sign that mountainous areas or larger ethnic minority populations had any impact on participation. A higher school enrollment rate, access to credit from the Agricultural Bank, a national road passing through the commune and distance to the province center had significant negative effects. More motorcycle-owning households and the presence of passenger transport in the commune improved the likelihood of being selected to participate in the project. Finally communes in Kon Tum

were also more likely to get a project than those in other provinces.

Figure 1 shows that there is imperfect overlap in the density of the estimated propensity scores for project and non-project communes. This suggests that it will be important to use propensity score matching in selecting comparison communes.

We use the resulting propensity scores to implement three matching estimators. In one we only trim non-project communes with p-scores lower than the lowest treatment's p-score. This results in a loss of 9 comparison communes. We refer to this as 'treatment support.' In the second matching method we further trim the project communes with p-scores that are higher than any of the non-project communes' thus limiting the sample to the region of common support. In addition to the 9 non-project observations, we lose 14 treatment communes. This provides tighter matching but also results in a loss of inference about the population of project villages since the sample is now changed. We report these estimates as a test of the sensitivity of the results.

Finally, the third matching procedure uses a caliper of 0.03 to discard comparison communes whose p-scores are further from treatment scores, and a nearest neighbor weighting scheme with replacement. In this case, we end up with 86 treatment communes matched with 38 comparison communes.

6.2 *Double difference results*

Table 5 presents unmatched and matched double difference estimates of the increment to rehabilitated road kilometers due to the project. These represent the before and after mean change in the road kilometers rehabilitated in the project communes over and above the before and after mean change in the non-project communes. We present double difference estimates for the entire project period, as well as for the intervening

mid-project situation in 1999. In the same way, double difference estimates of the increments to newly built road kilometers are also shown. Under our assumptions, these estimates reflect causal effects of the road rehabilitation project.

In addition, Table 5 gives the mean per commune kilometers rehabilitated by RTP1, as reported in the project data base. For our sample, RTP1 rehabilitated an average of just over four and a half kilometers per commune. Two years into implementation, the project had completed under half of that at a per commune average of 1.88 kilometers.

As noted earlier, a simple double difference comparison of the mean incremental kilometers rehabilitated by the 99 project and 100 non-project communes indicates a negligible and statistically insignificant difference of 0.08 kilometers by the end of the project period. The estimates suggest that midway into the project, project communes had actually rehabilitated 0.37 fewer kilometers than the comparison communes, although again, this estimate is not significantly different from zero. By contrast, the unmatched DD estimates indicate that the project communes actually built more new roads than the comparison communes. The difference is more pronounced and significant for the earlier period.

The estimates are qualitatively similar to the unmatched DD when we trim the comparison and/or treatment samples outside the region of common support (whether or not we trim the treatment group). There is little differential impact between treatment and non-treatment communes. These estimates suggest an extreme form of fungibility — whereby project funding was almost fully displaced to other uses — and no income effect on road spending. However, there does not appear to have been a diversion of project

funds as communes report having received international or central government funding and roads were indeed rehabilitated. These results point instead to a situation whereby the project paid for rehabilitation that the commune had already planned to do which allowed funding to be diverted to other uses, possibly including road improvements in non-project areas.

However, the picture changes substantially when we impose stronger matching criteria by using the caliper to exclude comparison communes with p-scores outside 0.03 of the score for a treatment commune, and weight the estimates thus tightening the matches. As can be seen in the last column of Table 5, the resulting DD estimate indicates a net gain of 3.06 kilometers attributable to the project. Although this estimate still reflects an average 33% displacement relative to the per commune average of 4.6 kilometers rehabilitated by RTP1, it indicates a markedly lower degree of displacement.

6.3 *Sensitivity test*

As discussed earlier, a number of communes may have reported rehabilitating roads when they meant maintenance. We have treated them as rehabilitation in the above estimates. We now test sensitivity to that choice by setting these rehabilitated kilometers to zero on the assumption that although roads were maintained, none were rehabilitated.

Table 6 presents the DD estimates under this assumption.

The differences between the unmatched and matched estimates, and between the alternative matched estimates, are now much smaller. The estimates range from the unmatched estimate of 2 kilometers rehabilitated over the life of the project to 2.9 kilometers under nearest neighbor matching. The latter indicates a displacement of 37 percent which is slightly higher than the equivalent estimate in Table 5.

A comparison of the estimates in Tables 5 and 6 suggest that tighter matching (column 3, Table 5) successfully deals with the problem we have identified in the data. The impact estimates increase with the weighted propensity score matching. It must then be the case that counterfactual observations that did less road rehabilitation work have initial level characteristics that are more like those of the treatment communes and so are better matches. Indeed, we find that of the 11 non-project communes that we classify as probable maintenance communes, only two end up included as matches in the weighted DD estimates. The others have low p-scores that are outside the range imposed by the caliper. This finding points to the importance of careful propensity score matching over and above using matching only to ensure common support.

Our results clearly indicate that the project has had impact on rehabilitated road kilometers, although not the full impact intended. There is at least some fungibility, assuming that the income effect of demand for roads is positive. However, given the possibility of an income effect, we cannot conclude that our results using tighter matching reject full fungibility.

6.4 *Impacts on road quality*

So far we have focused on total rehabilitated road kilometers without distinguishing between the types of roads that were rehabilitated. Nor have we addressed the issue of newly built roads. Yet, there is some indication in Tables 5 and 6 that the project also had a net impact on road building.

Table 7 presents a before and after project comparison of the breakdown in the types of roads that were rehabilitated and built in project communes. The underlying time periods differ but the percentage shares over time are comparable. Compared to

what the project areas were doing prior to the project, there are clear signs of a switch away from the rehabilitation of motor vehicle impassable earth roads and towards the rehabilitation of both paved all-weather, and paved sometimes impassable roads. Similarly, in building new roads, a compositional shift away from earth roads, particularly motor-vehicle impassable earth roads, towards paved roads is apparent relative to the pre-project period. Attribution of these shifts to the project is of course unclear without a comparison group. To examine this more rigorously, Table 8 presents double difference estimates of the increments of different road types attributable to the project. Here too, we see signs of a compositional shift whereby project communes focused their rehabilitation efforts on paved all weather, and paved sometimes impassable roads, more so than the non-project communes. However, in terms of road building, the impact of the project is felt primarily on the kilometers of earth all-weather roads and less so, although still significantly relative to non-project areas, on paved, sometimes impassable roads. Hence, the shift away from motor-impassable earth roads and towards paved all-weather roads seen for project communes over time in Table 7 also occurred in non-project communes.

These results strongly suggest that the project caused a shift in road rehabilitation efforts from earth roads to paved roads within project communes. Impacts of the project on the cost per kilometer rehabilitated provide additional support for this finding. The matched and weighted DD estimate of the project's impact on the per kilometer cost of rehabilitation shows an overall increase of 11.36 million dong in real 1995 prices, significant at the one percent level.¹² Furthermore, our results strongly suggest that part

¹² At the 1995 exchange rate of 11,000 dong to a dollar, this indicates an increase of \$1033 per kilometer.

of the displaced project funds were diverted to the building of new roads. Project communes ignored the project stipulation that they not build new roads. Through fungibility they imposed their own priorities and diverted some of the funding to priority roads. Our results clearly indicate that the project has had impact on road quality in the selected communes. This externally financed road project does appear to have achieved its immediate aims, though possibly at the cost of a switch away from earth road rehabilitation.

7. Conclusions

This paper has studied the issue of project aid fungibility in the context of a specific project using impact evaluation methods. We have estimated the impact of a World Bank-financed rural road rehabilitation project in Vietnam on the kilometers of roads actually rehabilitated.

A simple double difference estimate indicates that the project's net contribution to rehabilitated road increments is close to zero. This suggests large displacement of funding. On first inspection, the project funding appears to have gone entirely to uses other than road rehabilitation in project communes, possibly in part on road building. The rest may have resulted in lower taxes, or been spent on roads or other things in or outside the project communes.

However, this appears to reflect a bias in the simple difference-in-difference method. By better controlling for selection criteria and making the comparison communes look more like the treatment communes we find that what looks like fungibility is in fact attributable in large part to selection bias. In fact, the tighter form of matching appears to work in large part by correcting for measurement errors in our data.

When we implement this more rigorous impact evaluation methodology, we find much less evidence of displacement, though there is still some; we estimate that about one-third of the intended expansion to serviceable road length was displaced.

We also find that, in general, the quality of roads improved in the project communes. There is evidence of behavioral responses by implementing agents away from donor-imposed restrictions, notably in the evident switch in road rehabilitation from earth roads to paved roads as well as impacts on the kilometers of newly built roads.

A caveat on the interpretation of our results is that we cannot conclude that our commune level results reject the extreme fungibility model of external aid since we cannot rule out the possibility that the impact we do measure is entirely due to the income effect. Typically as communities get richer, roads improve. The income effect on the demand for roads can be assumed to be positive. Getting some displacement therefore implies some fungibility. But, we are unable to conclusively assess how much fungibility is present, given that we do not know the income effect of the project aid. A more complex evaluation design could in principle identify the income effect.

While impact evaluations have become more sophisticated in recent years — worrying about problems of selection and attrition bias, the heterogeneity of impacts and so on — it remains rare for evaluations to establish that an intervention actually funded what it intended and was supplemental to local spending. It is our conclusion from this analysis that this should be ascertained as a first step in impact evaluations. Although our results are less dramatic than they initially appeared, it will still be important to take the estimated funding displacement and deviations from project stipulations into account when attempting to determine impacts of the road rehabilitation on living standards.

References

- Dehejia, Rajeev H., and Sadek Wahba (1998), "Propensity Score Matching Methods for Non-Experimental Causal Studies," NBER Working Paper 6829, Cambridge, Mass.
- Dehejia, Rajeev H., and Sadek Wahba (1999), "Causal Effects in Non-Experimental Studies: Re-Evaluating the Evaluation of Training Programs," *Journal of the American Statistical Association*, 94, 1053-1062.
- Feyzioglu, Tarhan, Vinaya Swaroop and Min Zhu (1998), "A Panel Data Analysis of the Fungibility of Foreign Aid," *World Bank Economic Review* 12(1), 29-58.
- Gannon, Colin and Zhi Liu (1997). "Poverty and Transport." TWU discussion papers, TWU-30, World Bank, Washington, DC.
- Glewwe, Paul, Michele Gagnolati and Hassan Zaman Zaman (2002). "Who Gained from Vietnam's Boom in the 1990s?" *Economic Development and Cultural Change* 50(4), 773-92.
- Government of Vietnam–Donor Working Group. 2000. "Annex C: Commune-Level Fees and Contributions." In *Vietnam, Managing Resources Better: Public Expenditure Review 2000*. Vol. 2. *Joint Report on Public Expenditure Review*. Hanoi.
- Grootaert, Christiaan (2001), "Socioeconomic Impact Assessment of Rural Roads: Methodology and Questionnaires," mimeo, INFTD, World Bank.
- Grossman, Jean B. (1994), "Evaluating Social Policies: Principles and U.S. Experience." *The World Bank Research Observer* 9(2), 159-180.
- Heckman, James, H. Ichimura, and Petra Todd (1997), "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program," *Review of Economic Studies*, 64, 605-654.
- Heckman, James, H. Ichimura, Jeffrey Smith, and Petra Todd (1998), "Characterizing Selection Bias using Experimental Data," *Econometrica*, 66, 1017-1099.
- Heckman, James, H. Ichimura, Jeffrey Smith, and Petra Todd (1996), "Nonparametric characterization of selection bias using experimental data: A study of adult males in JTPA. Part II, Theory and Methods and Monte-Carlo Evidence," mimeo, University of Chicago.
- Jalan, Jyotsna and Martin Ravallion (1998), "Are There Dynamic Gains from a Poor-area Development Program?" *Journal of Public Economics* 67, 65-86.
- Jalan, Jyotsna and Martin Ravallion (2003), "Estimating the Benefit Incidence of an

- Antipoverty Program by Propensity Score Matching,” *Journal of Business & Economic Statistics* 21(1):19-30.
- Khilji, Nasir and Ernest Zampelli (1994), “The Fungibility of U.S. Military and Non-military Assistance and the Impacts on Expenditures of Major Aid Recipients,” *Journal of Development Economics* 43, 345-362.
- Pack, Howard and Janet Rothenberg Pack (1990), “Is Foreign Aid Fungible? The Case of Indonesia,” *The Economic Journal* 100, 188-194.
- Pack, Howard and Janet Rothenberg Pack (1993), “Foreign Aid and the Question of Fungibility,” *The Review of Economics and Statistics* 75(2), 258-265.
- Rosenbaum, Paul and Donald Rubin (1983), “The Central Role of the Propensity Score in Observational Studies for Causal Effects,” *Biometrika*, 70, 41-55.
- Rosenbaum, Paul and Donald Rubin (1985), “Constructing a Control Group using Multivariate Matched Sampling Methods that Incorporate the Propensity Score,” *American Statistician*, 39, 35-39.
- Rubin, Donald and Neal Thomas (2000), “Combining Propensity Score Matching with Additional Adjustments for Prognostic Covariates,” *Journal of the American Statistical Association* 95, 573-585.
- UNDP (2004), “Millenium Report: Interim Report of Task Force 1 on Poverty and Economic Development,” United Nations Development Program, New York.
- van de Walle, Dominique (1998). “Infrastructure and Poverty in Vietnam,” in Dollar, David, Paul Glewwe and Jennie Litvack (eds.) *Household Welfare and Vietnam’s Transition*, 99-135, World Bank Regional and Sectoral Studies, World Bank, Washington, D.C.
- van de Walle, Dominique (2005). “Evaluating the impacts of rural road projects,” mimeo, DECRG, World Bank, Washington D.C.
- World Bank (1996), “Rural Transport Project Staff Appraisal Report,” No. 15537-VN, East Asia and Pacific Region, World Bank, Washington, D.C.
- World Bank (1998), *Assessing Aid: What Works, What Doesn’t, and Why?* Oxford University Press: New York.
- Zampelli, Ernest M. (1986), “Resource Fungibility, the Flypaper Effect, and the Expenditure impact of Grants-in-Aid,” *Review of Economics and Statistics* 68(1), 33-40.

Table 1: Baseline average distances to closest geographical points from the commune center in project and non-project communes

	Project (N=100)	Non-project (N=100)
Big city (Hanoi, Haiphong, HCMC, Danang)	269.6 (146.0)	275.2 (130.3)
Provincial center	44.0** (36.5)	53.7 (38.3)
District center	13.5 (9.4)	13.1 (10.5)
National road ^a	9.5 (11.3)	9.2 (14.2)
if outside commune:	13.7* (11.2)	16.9 (15.6)
Provincial road ^a	6.3*** (9.0)	11.7 (18.2)
if outside commune:	10.2*** (9.6)	17.9 (19.9)
Railway station ^a	87.7 (100.3)	77.3 (86.6)
if outside commune:	104.8 (102.2)	85.9 (87.5)
River/canal port ^a	58.7 (91.6)	57.1 (69.5)
if outside commune :	79.7 (99.5)	72.2 (70.9)
Passenger transport	4.7 (8.6)	4.4 (8.2)

Notes:

Standard deviations in parentheses. ^a indicates that distances are averaged including zeroes for those communes that have the service within them. Standard deviations in parentheses. *, **, *** and **** indicates the difference across project and non-project communes are statistically significant at the 20, 10, 5 and 1% levels.

Table 2: Baseline mean data on the road situation in project and non-project communes

	Project (N=100)	Non-project (N=100)
% communes through which a national road passes	30.0*** (46.1)	44.0 (49.9)
km of national road in commune area (if in commune)	2.2 (4.6)	3.0 (4.7)
% communes through which a provincial road passes	38.0 (48.8)	32.0 (46.9)
km of provincial road in commune area (if in commune)	2.7* (4.9)	1.7 (3.2)
Total km of communal (rural) roads in commune:	32.7 (22.9)	28.8 (29.5)
<i>Paved all weather roads</i>	2.8 (9.1)	1.8 (3.6)
<i>Paved, sometimes impassable</i>	0.4 (1.5)	0.5 (1.9)
<i>Earth road, motor vehicle passable</i>	15.2 (12.4)	15.0 (18.9)
<i>Earth road, motor vehicle impassable</i>	14.3 (17.3)	11.5 (15.5)
% communes with rural roads built in last 3 years	53.0 (50.2)	47.0 (50.2)
km of new rural roads built in last 3 years:	4.5 (7.6)	3.6 (7.2)
<i>Paved, all weather roads</i>	0.2 (1.2)	0.3 (1.4)
<i>Paved, sometimes impassable</i>	0.03 (0.3)	0.1 (1.0)
<i>Earth road, motor vehicle passable</i>	3.2* (5.7)	2.1 (4.0)
<i>Earth road, motor vehicle impassable</i>	1.2 (3.9)	1.1 (4.6)
% communes with rural roads rehabilitated in last 3 years	78.0 (41.6)	77.0 (42.3)
km of rural roads rehabilitated in last 3 years:	11.4 (14.1)	10.2 (10.8)
<i>Paved, all weather roads</i>	0.6 (2.1)	0.5 (1.8)
<i>Paved, sometimes impassable</i>	0.1 (0.9)	0.2 (1.2)
<i>Earth road, motor vehicle passable</i>	6.2 (8.5)	6.4 (7.9)
<i>Earth road, motor vehicle impassable</i>	4.5 (9.3)	2.1 (4.1)
% communes with missing bridges	14.0 (34.9)	10.0 (30.2)

Notes: Standard deviations in parentheses. *, **, *** and **** indicates the difference across project and non-project communes are statistically significant at the 20, 10, 5 and 1% levels.

Table 3: Mean rural road kilometers by commune over project period

	1997		1999		2001	
	Project	Non-project	Project	Non-project	Project	Non-project
Total km of rural roads:	32.46	28.87	36.23*	30.78	38.06	34.59
Total built & rehabilitated last 2 years: <i>total km</i>	10.47	9.22	12.71	10.94	12.34	12.36
Built last 2 years: <i>total km</i>	2.99	2.39	4.45****	2.30	3.47	3.95
Rehabilitated last 2 years: <i>total km</i>	7.49	6.83	8.26	8.63	8.87	8.42

Notes:

*, **, *** and **** indicates single difference between project and non-project is significant at 20, 10, 5 and 1% levels; 199 observations.

Table 4: Logit regression of commune participation in RTP1

Variables	Coefficient	t-statistic
Province dummies: Tra Vinh	Reference	
Lao Cai	0.488	0.34
Thai Nguyen	1.992	1.20
Nghe An	1.915*	1.33
Binh Thuan	1.405	1.11
Kon Tum	3.602***	2.20
Terrain dummies: Coast	Reference	
Mountains	-0.953	-0.71
Upland	-0.877	-0.65
Plains	-1.070	-0.94
Population (thousands)	0.287****	2.89
Population density	-0.141	-1.28
log of predicted average per capita household expenditures	1.841*	1.45
Gross enrollment rate (age 6-15 enrolled over population age 6-15)	-5.356****	-3.24
Road density (kms of commune and district roads per 1000 hectares)	0.022*	1.44
Share of earth and car impassable roads in total km of roads	0.742	1.03
Commune performs routine road maintenance (dummy)	-0.064	-0.11
Passenger transport service is available in the commune (reference =1)	1.298***	2.13
Freight transport service is available in the commune	-0.116	-0.19
A national road passes through the commune	-2.375****	-4.00
A provincial road passes through the commune	-0.542	-1.18
A waterway passes through the commune	1.009*	1.51
A railway passes through the commune	1.019*	1.62
Distance from commune center to nearest big city (kms)	0.000	0.14
Distance from commune center to province center (kms)	-0.016**	-1.69
Distance from commune center to district center (kms)	0.015	0.59
Share of households with a Kinh ethnicity head	-0.991	-0.97
Share of households with a locally born head	0.797	0.88
Share of households with a pensioner	1.891	0.86
Share of households without land	0.333	0.14
Share of households with wage income	0.541	0.50
Share of households with cottage industries	1.498	0.96
Share of households in the trades or services	0.488	0.32
Commune has exploitable minerals (dummy)	-0.746	-1.14
Commune has a market (dummy)	-0.159	-0.35
Commune has an active land market (dummy)	0.359	0.70
Fulltime workers in government authorities per capita	-28.483	-0.61
Fulltime workers in mass organizations per capita	-164.083	-0.90
Fulltime workers in private enterprises per capita	161.891*	1.34
Share of households for whom farming is the main income	1.859	1.00
Unskilled nonagricultural employment is available in the commune (dummy)	0.120	0.27
Share of households with a permanent building as dwelling	0.495	0.59
Share of households owning motorcycles	5.496**	1.71
Agricultural Bank is available in the commune (dummy)	-1.385****	-2.57
Other sources of credit are available in the commune (dummy)	-0.006	-0.01
Government production organization is present in the commune (dummy)	-0.802*	-1.42
Constant	-12.886*	1.32
Log-likelihood	-103.845	
Pseudo R ²	0.251	
Number of observations	200	

Notes:

*, **, *** and **** indicates significance levels of 20, 10, 5 and 1%.

Table 5: Double difference estimates of rural road kilometer increments during project period (total km)

	Change between 1999 and 1997 Double difference				Change between 2001 and 1997 Double difference			
	Unmatched	(1) Treatment support	Matched (2) Common support	(3) Nearest neighbor	Unmatched	(1) Treatment support	Matched (2) Common support	(3) Nearest neighbor
<i>Estimated from SIRR:</i>								
Km of rural roads:								
Total built & rehabilitated:	1.77	1.79	1.30	4.76****	1.74	1.82	1.57	4.16**
Total built:	2.14***	2.28****	2.02***	3.20****	1.67*	1.63*	1.32	1.11
Total rehabilitated:	-0.37	-0.49	-0.72	1.57*	0.08	0.18	0.25	3.06**
<i>Data from project database:</i>								
Rehabilitated under RTP1:	1.88	1.88	1.76	1.82	4.58	4.58	4.64	4.54

Notes:

(i) Values for 2001 are the sum over 1999 and 2001.

(ii) *, **, *** and **** denote double difference between project and non-project over time is significant at 20, 10, 5, or 1% levels.

(iii) Samples are the following:

Unmatched: 99 treatment and 100 comparison communes

Matched: (1) 99 matched w/91 comparison

(2) 86 matched w/91 comparison

(3) 86 matched w/ 38 comparison

Table 6: Double difference estimates of rural road kilometer increments during project period (total km)

	Change between 1999 and 1997				Change between 2001 and 1997			
	Double difference				Double difference			
	Unmatched	(1) Treatment support	Matched (2) Common support	(3) Nearest neighbor	Unmatched	(1) Treatment support	Matched (2) Common support	(3) Nearest neighbor
setting observations to 0								
<i>Estimated from SIRR:</i>								
Km of rural roads:								
Total built & rehabilitated:	3.14****	3.29****	2.99***	4.70****	3.70***	3.79***	3.73***	4.01**
Total built:	2.15****	2.29****	2.03***	3.20****	1.69*	1.65*	1.35	1.13
Total rehabilitated:	0.99*	1.00*	0.97	1.49	2.01*	2.14*	2.38*	2.88*
<i>Data from project database:</i>								
Rehabilitated under RTP1:	1.89	1.89	1.76	1.83	4.62	4.62	4.69	4.60

Notes:

i) Values for 2001 are the sum over 1999 and 2001.

ii) *, **, *** and **** denote double difference between project and non-project over time is significant at 20, 10, 5, or 1% levels. The matched nearest neighbor sample is 86 communes matched with 38 comparisons.

Table 7: Shares of different road types built and rehabilitated in project communes

	Three years prior to project	During project
Average km of new rural roads built (km):	4.5	5.9
<i>with following road composition (%):</i>		
Paved, all weather roads	4.4	16.5
Paved, sometimes impassable	0.7	3.0
Earth road, motor vehicle passable	71.0	62.4
Earth road, motor vehicle impassable	26.7	18.1
Average km of rural roads rehabilitated:	11.4	13.1
<i>with following road composition (%):</i>		
Paved, all weather roads	5.3	20.3
Paved, sometimes impassable	0.9	4.9
Earth road, motor vehicle passable	54.4	57.9
Earth road, motor vehicle impassable	39.5	17.0

Table 8: Double difference estimates of increments to different types of rural roads during project period (total km)

	Change between 2001 and 1997			
	Double difference			
	Unmatched	(1) Treatment support	(2) Common support	(3) Nearest neighbor
<i>(km of rural roads)</i>				
Total built:	1.67*	1.63*	1.32	1.11
<i>Of which:</i>				
Paved, all-weather	-0.53	-0.67	-0.69	-0.27
Paved, sometimes impassable	0.14*	0.13*	0.16*	0.19**
Earth, all-weather	1.59***	1.58***	1.03	2.06****
Earth, sometimes impassable	-0.22	0.04	-0.02	0.34
Total rehabilitated:	0.08	0.18	0.25	3.06**
<i>Of which:</i>				
Paved, all-weather	1.08***	1.01**	0.96*	1.36***
Paved, sometimes impassable	0.45**	0.49***	0.47**	0.56***
Earth, all-weather	0.73	1.05	0.91	1.91
Earth, sometimes impassable	-0.86	-0.75	-0.64	0.21

Notes:

(i) Values for 2001 are the sum over 1999 and 2001.

(ii) *, **, *** and **** denote double difference between project and non-project over time is significant at 20, 10, 5, or 1% levels.

(iii) Samples are the following:

Unmatched: 99 treatment and 100 comparison communes

Matched: (1) 99 matched w/91 comparison

(2) 86 matched w/91 comparison

(3) 86 matched w/ 38 comparison

Figure 1: Density of estimated propensity scores of project and non-project communes

